

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

See included English Abstracts
and 2 seq. search reports

(19) 日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平9-121870

(43) 公開日 平成9年(1997)5月13日

(51) Int.Cl. ⁶	識別記号	序内整理番号	F I	技術表示箇所
C 12 N 15/09	ZNA	9162-4B	C 12 N 15/00	ZNAA
1/19			1/19	
1/21			1/21	
9/50			9/50	
// (C 12 N 1/19				

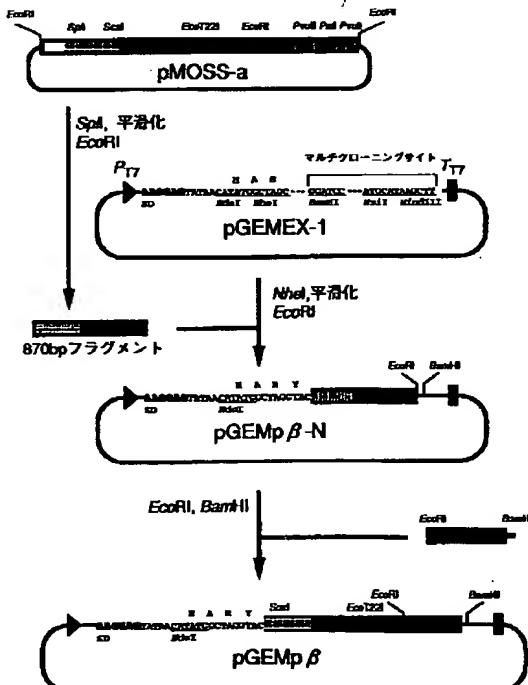
審査請求 未請求 請求項の数18 FD (全17頁) 最終頁に続く

(21) 出願番号	特願平7-353931	(71) 出願人	000000066 味の素株式会社 東京都中央区京橋1丁目15番1号
(22) 出願日	平成7年(1995)12月28日	(72) 発明者	河合 美佐子 神奈川県川崎市川崎区鈴木町1番1号 味の素株式会社食品総合研究所内
(31) 優先権主張番号	特願平6-340399	(72) 発明者	浅野 皆夫 神奈川県川崎市川崎区鈴木町1番1号 味の素株式会社食品総合研究所内
(32) 優先日	平6(1994)12月29日	(72) 発明者	鈴木 俊一 神奈川県川崎市川崎区鈴木町1番1号 味の素株式会社中央研究所内
(33) 優先権主張国	日本 (JP)	(74) 代理人	弁理士 長谷川 洋子 (外2名) 最終頁に続く
(31) 優先権主張番号	特願平7-245279		
(32) 優先日	平7(1995)8月30日		
(33) 優先権主張国	日本 (JP)		

(54) 【発明の名称】 新規チオールプロテアーゼをコードするDNAおよびこれを用いた該チオールプロテアーゼの製造方法

(57) 【要約】

【解決手段】 発芽ダイズ子葉に由来する新規チオールプロテアーゼをコードするcDNA、該DNAを含有する組換え発現ベクター、該組換え発現ベクターによって形質転換された大腸菌、酵母である形質転換体並びに該形質転換体を用いたチオールプロテアーゼの製造方法。
【効果】 ダイズタンパク質を効率よくアミノ酸または低分子ペプチドにまで分解し得る新規チオールプロテアーゼを、遺伝子組換え手法により直接的にしかも高効率で大量に生産することが可能となる。



1

【特許請求の範囲】

【請求項1】 配列表の配列番号1または2に示す塩基配列で表される、発芽ダイズ子葉に由来する新規チオールプロテアーゼをコードするDNA。

【請求項2】 配列表の配列番号1または2におけるアミノ酸配列またはこの配列において1若しくは複数のアミノ酸残基が挿入、付加、欠失若しくは置換されたアミノ酸配列を有し、発芽ダイズ子葉に由来する天然型チオールプロテアーゼの活性特性を発揮し得るチオールプロテアーゼをコードするDNA。

【請求項3】 配列表の配列番号3に示す塩基配列またはその部分配列で表されるDNAであって、該配列番号3におけるアミノ酸配列またはこの配列において1若しくは複数のアミノ酸残基が挿入、付加、欠失若しくは置換されたアミノ酸配列を有し、発芽ダイズ子葉に由来する天然型チオールプロテアーゼの活性特性を発揮し得るチオールプロテアーゼをコードする、上記DNA。

【請求項4】 配列表の配列番号3のアミノ酸番号4位(Tyr)～351位(Va1)までのアミノ酸配列をコードする、請求項3に記載のDNA。

【請求項5】 配列表の配列番号3の塩基番号10～1053の塩基配列で表される、請求項3または4に記載のDNA。

【請求項6】 配列表の配列番号3のアミノ酸番号111位(Asp)～351位(Va1)までのアミノ酸配列をコードする、請求項3に記載のDNA。

【請求項7】 配列表の配列番号3の塩基番号331～1053の塩基配列で表される、請求項3または6に記載のDNA。

【請求項8】 請求項1～7のいずれか1項に記載のDNAを含有する、組換え発現ベクター。

【請求項9】 請求項8に記載の組換え発現ベクターで形質転換された大腸菌(*Escherichia coli*)である、形質転換体。

【請求項10】 請求項9に記載の形質転換体を培養し、培養物中に產生されたチオールプロテアーゼを単離し、所望により精製することからなる、チオールプロテアーゼの製造方法。

【請求項11】 前記培養物中に產生されたチオールプロテアーゼをタンパク質封入体として回収し、可溶化・巻き戻した後にpHを酸性側にシフトして活性化することからなる、請求項10に記載のチオールプロテアーゼの製造方法。

【請求項12】 請求項4または5に記載のDNAの5'末端側にシグナル配列をコードするDNAを付加してなるDNA。

【請求項13】 前記シグナル配列が、配列表の配列番号2の塩基番号48～122の塩基配列である、請求項12に記載のDNA。

【請求項14】 配列表の配列番号2のアミノ酸番号一

2

132位(Met)～241位(Va1)までのアミノ酸配列をコードするDNA。

【請求項15】 配列表の配列番号2の塩基番号48～1166の塩基配列で表されるDNA。

【請求項16】 請求項12～15のいずれか1項に記載のDNAを含有する、組換え分泌発現ベクター。

【請求項17】 請求項16に記載の組換え分泌発現ベクターで形質転換された酵母(*Saccharomyces cerevisiae*)である、形質転換体。

10 【請求項18】 請求項17に記載の形質転換体を培養し、培地中に產生されたチオールプロテアーゼを単離し、所望により精製することからなる、チオールプロテアーゼの製造方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、従来のプロテアーゼでは分解の難しいダイズ種子貯蔵タンパク質を容易に分解し得る新規なチオールプロテアーゼをコードするDNA、該DNAを含有する組換え発現ベクター、該組換え発現ベクターにより形質転換された形質転換体並びに該形質転換体を用いたチオールプロテアーゼの製造方法に関する。

【0002】

【従来の技術】ダイズタンパク質のアミノ酸への分解は、塩酸等の酸による加水分解、あるいは麹菌等の微生物酵素をはじめとする既存のプロテアーゼによる分解などにより行われている。

30 【0003】しかしながら、酸による加水分解を行った場合、天然調味料になり得るようなダイズタンパク質加水分解産物を得ようとすると、100°Cで1～2日間の反応が必要となり、このような高温、長時間の反応はエネルギー消費量が大きいという問題がある。さらに、酸によるタンパク質の加水分解法は簡便である一方、アミノ酸の過剰分解(破壊)、中和のために高塩分となるなどの問題もある。

【0004】そこで上記問題を解決すべく、既存のプロテアーゼによる穏やかな条件下でのタンパク質分解が提案され、特に、タンパク質分解酵素(プロテアーゼ)によるダイズタンパク質のアミノ酸への分解は、生物反応を利用する穏やかな反応条件で進行するため、化学反応を利用する酸加水分解に代わる方法として実現が期待されている。

40 【0005】しかしながら、一般にマメ科植物の貯蔵タンパク質は、未変性の状態では既存のプロテアーゼに対してかなりの耐性を有することが知られており(S. S. Nielsen et al., "J. Agric. Food Chem.", 36, 896(1988))、これらを酵素的に完全分解に至らしめるには厳密な条件下での変性処理が必要とされ、そのうえ、反応には長時間を要する。この長時間の反応は、反応基質の完全な滅菌が困難なことと相まって、反応中の雑菌汚染

50

の原因となる。そのため、常に高濃度の塩や酢酸等の添加が必要になるという問題がある。

【0006】そのためダイズタンパク質を変性処理の有無にかかわらず短時間で容易にアミノ酸または低分子ペプチドにまで分解し、また、雑菌の混入しにくい酸性やアルカリ性域でも反応し得るような酵素の発見が望まれていた。

【0007】本発明者らは、このような期待に応え得るプロテアーゼを発芽ダイズ子葉より見出し、単離精製を行い、諸性質を調べ、そのN末端付近のアミノ酸配列を決定し、さらに酵素活性を確認し、従前に出願を行っている（特願平6-294548号明細書）。すなわち、当該プロテアーゼはプロテアーゼD3とも名付けられた新規チオールプロテアーゼであり、分子量はSDS-PAGEで26~30kDa、反応至適pH約3~7、至適温度30~50°Cで、ウシ血清アルブミンのみならず、ダイズタンパク質であるグリシン（11Sグロブリン）、β-コングリシン（7Sグロブリン）を未変性状態でアミノ酸、あるいはオリゴペプチドまできわめて強力に分解し得る特異な酵素であり、その産業上の利用が大いに期待される。

【0008】しかしながら、この新規チオールプロテアーゼの発芽ダイズ子葉中の含量は微量であるため、当該酵素を発芽ダイズ子葉抽出物より大量に調製することには相当な困難があった。

【0009】

【発明が解決しようとする課題】上述の問題点を解決する方法の一つとしては、該新規チオールプロテアーゼ遺伝子をダイズ以外の系で遺伝子組換え技術を用いて大量発現させることにより、大量のチオールプロテアーゼを取得するという方法が挙げられる。この方法を実現するためには、まず、当該新規チオールプロテアーゼをコードするcDNAを取得し、塩基配列を解析し、該チオールプロテアーゼの全アミノ酸配列に関する情報を得ることが必須である。

【0010】次に、該新規チオールプロテアーゼをコードするDNAを適当な発現ベクターに組み込み、目的物を大量に生産する形質転換体を得ることが必須である。

【0011】さらに場合によっては、この形質転換体の培養によって得られた該チオールプロテアーゼ遺伝子産物を、活性体へと変換する操作技術も必要であろう。

【0012】本発明はこのような事情に鑑みてなされたもので、発芽ダイズ子葉に由来する新規チオールプロテアーゼを、天然から単離するという方法に代えて、遺伝子組換え技術を用いて効率的な遺伝子発現並びに該チオールプロテアーゼの大量生産をなし得るための技術を提供することを課題とする。

【0013】

【課題を解決するための手段】本発明者らは上記課題を解決するために鋭意研究を行い、すでに本発明者らによ

り明らかにされている当該新規チオールプロテアーゼのN末端アミノ酸配列をもとに、発芽ダイズ子葉mRNAより調製したcDNAライブラリーをスクリーニングし、該チオールプロテアーゼをコードするcDNAを取得した。

【0014】本発明者らはさらに研究を進め、この得られたcDNAを用いて該新規チオールプロテアーゼを生産する形質転換体である大腸菌および酵母を得、これにより該チオールプロテアーゼ活性を有するタンパク質を得るに至った。

【0015】すなわち本発明は、発芽ダイズ子葉に由来する新規チオールプロテアーゼをコードするcDNA、および該cDNAを用いた遺伝子組換え技術による該チオールプロテアーゼの製造方法に関する。

【0016】

【発明の実施の形態】以下、本発明について詳述する。上述したように、発芽ダイズ子葉に由来する新規チオールプロテアーゼは、本発明者らにより単離に成功し、その特性についても研究されている（特願平6-294548号明細書）。この新規チオールプロテアーゼには2種類のアイソザイムが存在し、いずれも分子量、反応至適pH、至適温度、阻害剤の種類、活性化剤の種類、ダイズ種子貯蔵タンパク質をアミノ酸または低分子ペプチドにまで分解するという作用のいずれにおいても著しい一致をみている。したがって、本発明においてはいずれのアイソザイムも本発明でいう新規チオールプロテアーゼに含むものとする。

【0017】なお、上記2種類のアイソザイムを区別するため、便宜上一方をD3- α 、他の方をD3- β と称する。本発明者らにより既に確認されているD3- α のN末端アミノ酸配列は、Asp-Lys-Leu-Pro-Glu-Ser-Val-Asp-Trp-Arg-Lys-Glu-Gly-Ala-Val-Pro-Pro-Val-Lys-Asp-Gln-Gly-Gly-Xaa-Gly-Ser-Xaa-Trp-Ala-Pheであり、D3- β のN末端アミノ酸配列は、Asp-Lys-Leu-Pro-Asp-Ser-Val-Asp-Trp-Arg-Lys-Glu-Gly-Ala-Val-Pro-Pro-Val-Lys-Asp-Gln-Gly-Glyである。D3- α 中のXaaは未同定のアミノ酸であるが、これらはCysと推定される。両者のアミノ酸配列は、この範囲では5番目の残基のみが異なっていた（D3- α ではGlu、D3- β ではAsp）。なお、これら新規プロテアーゼのアイソザイムD3- α とD3- β を、D3と総称する。

【0018】本発明のcDNAの取得方法の具体例としては、例えば、D3のN末端のアミノ酸配列をもとにcDNAプローブを作製し、発芽ダイズ子葉から抽出したD3のmRNAを錆型に作製したcDNAライブラリーからハイブリダイゼーションによりD3のcDNAを釣り上げるなど、従来より慣用的に行われている方法が挙げられる。

【0019】D3のmRNA抽出に用いる発芽ダイズ子葉は、そのダイズの種類を問わない。すなわち市販され

ているダイズ、搾油原料として用いられているダイズ等、その栽培産地、品種を限定しない。また、発芽の方法、栽培条件、発芽の有無、発芽後の期間を問わなないが、ダイズ種子を吸水させ、10日間生長させた発芽ダイズ子葉を用いるのが好ましい。

【0020】また、cDNAプローブは、D3のN末端アミノ酸配列をもとに化学合成してもよく、あるいはPCR法等により作製してもよい。本実施例においてはPCR法により作製したが、その詳細については後述の実施例に詳しく記す。なお、cDNAライブラリーの作製方法は、常法により行うことができる。

【0021】次に、このようにして得たcDNAを発現ベクターに組み込んで組換え発現ベクターを作製する。用いるベクターは、特に限定されるものではないが、宿主細胞内で自律的に複製可能であって、上記DNAすなわちD3遺伝子を組み込み得る挿入部位をもち、さらにこの組み込んだDNAを宿主細胞内で発現せしめることを可能とする領域を有することが必要である。

【0022】なお、ベクターに組み込むD3遺伝子としては、cDNAだけでなく、cDNA配列から予測されるD3のアミノ酸配列をコードするよう設計して合成されたDNAでもよい。このようなアミノ酸配列をもとにした遺伝子の合成は、DNA自動合成機を利用して合成したオリゴヌクレオチドをアニール後に連結すること等により、容易に行うことができる。

【0023】さらに、D3- α とD3- β 間ではアミノ酸置換がみられても酵素的性質は著しく一致していることからも明白なように、一部のアミノ酸残基が置換しても、天然のD3と同じ酵素的性質が得られることが予測される。また、末端の残基が付加または欠失したり、D3とは関連のない配列、例えば発現プラスミド構築の際のリンカーDNAの配列などが付加したりしていても、結果として天然のD3と同じ酵素的性質が得られる可能性もある。このような遺伝子の変化は、市販の遺伝子の部位特異変異導入キットを用いたり、合成遺伝子を挿入することなどにより、容易に実現し得る。したがって、このベクターに組み込むD3遺伝子としては、天然のD3と同じ酵素的性質が維持される限り、その変異体もここでいうD3遺伝子に含む。

【0024】また、D3の生産方法として、異種タンパク質との融合タンパク質として生産させる方法もある。これは例えばpGEXシステム（ファルマシア社製）などを用いればD3をグルタチオン-S-トランスフェラーゼとの融合タンパク質として大腸菌で生産させることができる。

【0025】そのほかのD3の生産方法として、枯草菌、酵母、麹菌などのタンパク質分泌能を利用して、培地中へ生産させる方法もある。例えば酵母を宿主とした場合、D3自身のシグナル配列や、酵母自身の分泌タンパク質のシグナル配列、たとえば α ファクターなどのシ

グナル配列を用いれば、D3を培地中に分泌生産させることも可能である。

【0026】次いで上記組換え発現ベクターを宿主細胞内に導入し、形質転換体を得る。組換え発現ベクターの宿主細胞内への導入方法は、従来より慣用的に用いられている方法により行うことができる。コンピテントセル法、プロトプラスト法、リン酸カルシウム共沈法、エレクトロポレーション法、マイクロインジェクション法、リボソーム融合法等、種々のものが挙げられるが、用いる宿主に応じてそれぞれ任意の方法を取り得る。本発明のD3を产生する宿主としては、大腸菌、枯草菌、酵母、麹菌等の微生物のほか、カイコ培養細胞等の細胞を用いてもよい。

【0027】そして、このようにして得られた形質転換体を培養することにより、培養物中にD3を產生させる。これを公知の方法で単離し、場合によっては精製することにより、目的とする酵素が得られる。

【0028】形質転換体を培養するための培地は公知であり、例えば大腸菌ではLB培地などの栄養培地や、M9培地などの最小培地に炭素源、窒素源、ビタミン類などを添加して用いることができ、酵母ではYPDなどの栄養培地や、SD培地などの最小培地に炭素源、窒素源などを添加して用いることができる。形質転換体の培養は、宿主に応じて、通常16~42℃、好ましくは25~37℃で5~168時間、好ましくは8~72時間行う。振盪培養と静置培養のいずれも可能であるが、必要に応じて攪拌や通気を行ってもよい。また、D3遺伝子発現のために誘導型プロモーターを用いた場合は、培地にプロモーター誘導剤を添加して培養を行う。

【0029】D3の単離・精製法としては、形質転換体の抽出物や培地より、公知の塩析、等電点沈澱法、または溶媒沈澱法等の沈澱法、透析、限外済過またはゲル済過等の分子量の差を利用する方法、イオン交換クロマトグラフィー等の荷電の差を利用する方法、アフィニティーコロマトグラフィー等の特異的親和性を利用する方法、疎水クロマトグラフィー、逆相クロマトグラフィー等の疎水度の差を利用する方法等が挙げられ、これらを組み合わせることにより単離・精製が可能である。

【0030】大腸菌等を宿主とした場合には、不活性なD3会合体、すなわちタンパク質封入体としてD3遺伝子産物を得た後、これを適当な方法で活性化するのが好ましい。タンパク質封入体としてD3遺伝子を発現させた場合の利点として、D3遺伝子産物を封入体として簡単に単離できる、D3活性の宿主への影響を軽減できる、等の点が挙げられる。後述の実施例では、大腸菌を宿主として、一部改変したD3遺伝子をタンパク質封入体として高発現させて回収し、可溶化・巻き戻しした後にpHを酸性側にシフトして活性化させて本チオールプロテアーゼを得る方法について詳しく述べる。なお、タンパク質封入体を可溶化・巻き戻しした後にpHを酸性

側にシフトして活性化させる方法は公知であり、例えばキモシンに関しては "Biotechnology", September 1984, pp. 801-804に、カテプシン-Lに関しては "The Journal of Biological Chemistry", vol. 264, No. 34, pp. 20487-20495 (1989) に、それぞれ記載されている。すなわち、キモシン、カテプシン-Lは、pHを酸性にシフトさせることにより、成熟型酵素のN末端側にある配列を、酸性で機能するプロテアーゼによる消化や、自己消化により除去している。しかしながら、本チオールプロテアーゼD3に関しては、cDNA構造よりそのC末端側にも天然型D3にはない配列の存在が示唆された。そのため、あらかじめこのC末端側の配列を除いた形でD3を発現させ、その封入体について、可溶化・巻戻しを行った後にpHを酸性側にシフトさせて活性化を行った。

【0031】酵母や麹菌等、タンパク質分泌能を有する宿主を用いた場合には、D3を培地に分泌生産すると、D3の折り畳みが正常に行われるためD3活性体を得られるため、D3活性の宿主への影響が軽減でき、さらに精製が簡便になる等の利点がある。後述の実施例では、酵母を宿主として、一部改変したD3遺伝子を用いてD3を分泌発現させることにより、本チロールプロテアーゼを得る方法について詳しく述べる。

【0032】

【実施例】以下、本発明を実施例に基づいて説明する。

【0033】【実施例1 PCR法によるチオールプロテアーゼD3のN末端アミノ酸をコードする部分cDNA断片のクローニング】cDNAライブラリーのスクリーニングのためのプローブに用いるために、プロテアーゼD3のN末端部分をコードする部分cDNA断片を下記に示すようにPCR法により増幅後、クローニングした。

【0034】まず、播種後10日目のダイズ子葉を用いて、常法により、RNAエクストラクションキット(RNA Extraction Kit; ファルマシア社製)によりRNAを抽出し、mRNA精製キット(mRNA Purification Kit; ファルマシア社製)を用いて、poly(A) RNAを精製した。

【0035】このようにして得たpoly(A) RNAを錠型にして、いわゆるRT-PCR反応(reverse transcript polymerase chain reaction)を行ったのであるが、このときに用いたPCRプライマーは次のようなものであった。すなわち、本発明者らによりすでに決定されている上記D3- α のN末端配列中、1番目のAspから5番目のGluまでの配列に基づいて、5'-GA(C/T)AA(A/G)TT(A/G)CC(A/G/C/T)GA-3'なる縮重したオリゴヌクレオチドを合成し、これをセンス鎖のPCRプライマーとした。

【0036】また、26番目のSerから30番目のPheまでの配列(未同定のアミノ酸は他のチオールプロ

テアーゼとの比較よりCysと推定した)に基づいて、5'-AA(A/G/C/T)GCCCA(C/T)CA(C/T)CT-3'なる同様に縮重したオリゴヌクレオチドを合成し、これをアンチセンス鎖のPCRプライマーとした。

【0037】次に、これらのpoly(A) RNA、PCRプライマーを用いて、RT-PCR反応を行った。具体的には、poly(A) RNA 20ngを錠型として、RNA-PCRキット(パーキン・エルマー・シータス社製)を用いてoligo(dT)プライマーにより最初の1本鎖cDNAを合成した後、上記センス鎖プライマーおよびアンチセンス鎖プライマーをそれぞれ1μM存在下、DNAサーマルサイクラー(パーキン・エルマー・シータス社製)を用いて以下の温度条件でPCR反応を行った。

【0038】すなわち、94℃、5分間の熱変性の後、37℃、2分間のアニーリングを行い、続いて、1分30秒間の勾配を掛けて72℃まで温度上昇させた後、さらに1分30秒間の伸長反応を行った。この後、94℃、1分間の熱変性、37℃、2分間のアニーリング、1分30秒間の勾配を掛けた72℃までの温度上昇、72℃、1分30秒間の伸長反応からなるサイクルを35サイクル繰り返し、その後、72℃、7分間の伸長反応を行って反応を停止した。

【0039】この結果、約90bpの增幅断片が得られたので、これを、TAクローニングキット(インビトロゲン社製)によりプラスミドベクター(PCR2)にクローニングし、その塩基配列をALF DNAシークエンサー(ファルマシア社製)により決定した。

【0040】その結果、得られた塩基配列のコードするアミノ酸配列は、決定されたチオールプロテアーゼD3- α のN末端配列と完全に一致した。

【0041】【実施例2 チオールプロテアーゼD3 cDNAのクローニング】実施例1で得られた部分cDNA断片をプローブにしてcDNAライブラリーをスクリーニングし、チオールプロテアーゼD3のcDNAのほぼ全長を下記の方法によりクローニングした。

【0042】すなわち、実施例1での方法と同様にして得た発芽10日のダイズ子葉から抽出したpoly(A) RNA 5μgから、タイムセーバー cDNAクローニングキット(アマーシャム社製)を用いて2本鎖cDNAを合成し、これをcDNAラピッドクローニングシステム(アマーシャム社製)を用いて、常法により、λMOSS10×ベクターにクローニングし、cDNAライブラリーを得た。

【0043】このcDNAライブラリーの50,000pfuを実施例1で得た90bpの部分cDNA断片をプローブにして、ECLダイレクトDNA/RNAラベリング検出システム(アマーシャム社製)を用いて、ブラークハイブリダイゼーションによりスクリーニングした。

【0044】その結果、約50個の陽性クローンを得た。この中より20個をランダムにピックアップし、ブラークを純化した後、cDNAを入MOSS1oxベクターのもつ自動切り出し機能を利用して、プラスミドの形で回収した。

【0045】これらのプラスミドを単離し、90bp断片をプローブとしたサザンハイブリダイゼーション分析を行い、プローブとハイブリダイズするcDNAインサートをもつクローンをさらにピックアップした。

【0046】得られたクローンのcDNAインサートを制限酵素EcoRI、BglIII、PstIにより切断し、その切断パターンをアガロースゲル電気泳動により分析したところ、2種類のパターンに分れ、大部分のクローンはこのいずれか一方のパターンのグループに属していた。それぞれのグループから最大のcDNAインサートをもつクローンpMOSS-a (cDNAの長さは約1.9kb) およびpMOSS-r (cDNAの長さは約1.8kb) を選んで、その塩基配列を分析した結果、pMOSS-aにはチオールプロテアーゼD3-βのN末端アミノ酸配列をコードする部分が含まれており、pMOSS-rにはチオールプロテアーゼD3-αのN末端アミノ酸配列をコードする部分が含まれていることが明らかになった。

【0047】なお、D3-α cDNAは、pUC18プラスミドベクターに組み込まれた形で大腸菌(Escherichia coli) HB101に形質転換され、かかるプラスミド保持菌は、Escherichia coli AJ13067として、工業技術院生命工学工業技術研究所に平成6年12月6日に寄託されている(受託番号: FERM P-14688)。

【0048】また、D3-β cDNAは、pUC18プラスミドベクターに組み込まれた形で大腸菌(Escherichia coli) HB101に形質転換され、かかるプラスミド保持菌は、Escherichia coli AJ13066として、工業技術院生命工学工業技術研究所に平成6年12月6日に寄託されている(受託番号: FERM P-14687)。

【0049】これらのcDNAの全塩基配列を決定した。その結果、チオールプロテアーゼD3-αのcDNAの塩基配列は配列表の配列番号1に、チオールプロテアーゼD3-βのcDNAの塩基配列は配列表の配列番号2に記載した通りであった。

【0050】また、決定したN末端アミノ酸配列を含むフレームで上記塩基配列を翻訳して得られるアミノ酸配列は、配列番号1、2に示したとおりであった。

【0051】[実施例3 D3-βの大腸菌での生産] 実施例2で得られたD3-β cDNAの一部を大腸菌で機能する発現ベクターに組み込み、この発現プラスミドを保持する形質転換体を培養し、遺伝子産物をタンパク質封入体として得た。この封入体を菌体から取り出

し、in vitro で可溶化・巻き戻しを行った後、pHを酸性側にシフトさせると、初めてD3活性を得た。

【0052】大腸菌で生化学的に不活性なタンパク質封入体としてD3遺伝子を発現させると、D3遺伝子産物を封入体として簡便に単離でき、またD3活性の宿主菌体への影響を軽減できるなどの利点がある。

【0053】このときベクターに組み込んだcDNAは、配列表の配列番号3のアミノ酸配列番号4~351位をコードする部分(すなわち塩基番号10~1053部分)である。これは、配列番号2に示されるアミノ酸配列番号-107~241位をコードする部分(すなわち塩基番号123~1166部分)にあたる。すなわちこれは、D3-β cDNAの塩基配列より推定される全一次構造から、N末端側のアミノ酸配列番号-132位のMet~-108位のSer部分、およびC末端側のアミノ酸配列番号242位のCys~332位のA1a部分を除いたものである。

【0054】D3-βのこの部分配列を発現させるよう選択したのは、次の理由による。N末端については、プロテアーゼの成熟体のN末端側に存在する配列、いわゆるプロ配列が折り畳みおよび活性制御に関与することが知られており、D3-βではこのプロ配列は配列表の配列番号2のアミノ酸配列番号-107位近傍から-1位部分であると考えられたからである。

【0055】また、C末端については、ペプチドマッピングにより、C末端の位置にはヘテロジエニティがあり、配列表の配列番号2のアミノ酸配列番号228位のPro~247位のSerであることが判明した。しかし、他のチオールプロテアーゼにはD3-βのC末端側のアミノ酸配列番号242位のCys近傍~332位のA1a部分に相当する配列が存在しないこと、D3の分子量は26~30kDaであること、他のチオールプロテアーゼとの比較から242位のCysは活性への関与もS-S結合への関与もないと推測されること、242位のCysコドン(TGT)は1塩基置換により終止コドンになり得ること、等から、上述のようにC末端の位置を241位のValとした。

【0056】また、D3-αとD3-βの酵素化学的性質はほとんど一致しており、実施例2で既述したようにcDNA構造から両者のアミノ酸配列もほとんど一致していると考えられる。したがって、D3-αも本実施例で述べるD3-βの生産方法と同様な方法で生産することが可能であることが容易に推察できる。

【0057】以下に大腸菌を用いた本実施例について、<1>発現プラスミドの構築、<2>該発現プラスミドを用いた大腸菌の形質転換体の作製と培養、<3>タンパク質封入体の単離、可溶化・巻き戻し、<4>自己触媒的活性化について分説する。

【0058】<1>発現プラスミドの構築

50 図1に示すように、実施例2に記載のD3-β cDN

11

Aを組み込んだpMOSS-aを制限酵素S p I Iで切断し、K l e n o w 酵素で平滑末端化した後、E c o R Iで切断して約870bpの断片を得た。この断片を、市販の発現ベクターpGEMEX-1（プロメガ社製）をN h e Iで切断、K l e n o w 酵素で平滑末端化した後にE c o R Iで切断した大断片と連結し、プラスミドpGEMp β -Nを作製した。

【0059】一方、図2に示すように、pMOSS-aをE c o R I、P v u I Iで切断して得られた190bpの断片をM13mp19のE c o R I、S m a I間に挿入し、M13 β -Cを構築した。そして、M13 β -Cの一本鎖DNAを錆型に、オリゴヌクレオチド5'-CA AATGTCCTGAGACAACTACT-3'を変異プライマーとして、スカルプター インビトロ ミュータジェネシス システム（アマーシャム社製）を用い、TGT（配列番号2のアミノ酸配列番号242位のCysコドン）をTGA（終止コドン）に部位特異的に変異させ、M13 β -C(C242Opa1)を作製した。次いでこれをE c o R I、B a m H Iで切断して約190bpの断片を得た。

【0060】その後、図1に示すように、pGEMp β -NをE c o R I、B a m H Iで切断して得られる大断片と、上記M13 β -C(C242Opa1)からの約190bpの断片とを連結し、発現プラスミドpGEMp β を構築した。すなわちこのpGEMp β に発現するよう組み込んだ配列は配列表の配列番号3に記載したものであり、このなかで塩基番号1~9の配列はベクター由来の塩基配列である。

【0061】以上の変異や構築は、A L F DNA シークエンサー（ファルマシア社製）を用いて塩基配列を分析することにより正しく行われていることを確認した。

【0062】<2>発現プラスミドpGEMp β を用いた大腸菌の形質転換体の作製と培養
コンピテントセル法により、上記のpGEMp β を大腸菌J M 109 (DE3) 株（プロメガ社製）に導入した。得られた形質転換体を100 μ g/m lのアンピシリンを含むM9-カザミノ酸培地で37°Cで振盪培養した。培地にイソプロピル- β -D-チオガラクトピラノシドを添加して生産誘導すると、目的のD3遺伝子産物は、タンパク質封入体として菌体内に蓄積した。

【0063】<3>タンパク質封入体の単離・可溶化・巻き戻し

上述のようにして培養した菌体を集菌後、超音波破碎し、遠心によってタンパク質封入体を回収した。このタンパク質封入体を洗浄後、8M尿素-10mMジオトレイトール-50mM塩化ナトリウム-50mMトリス・塩酸-5mMエチレンジアミン四酢酸溶液(pH8)に、タンパク質濃度が約10mg/m lになるよう溶解した。これを可溶化p D 3- β 溶液と呼ぶ。

12

【0064】可溶化p D 3- β 溶液1容に、還元型グルタチオンおよび酸化型グルタチオンを含み、タンパク質変性剤を含まない溶液100容をゆっくり添加して、尿素濃度を低下させるとともに、タンパク質濃度も低下させた。この溶液の組成は1mM還元型グルタチオン-0.1~0.5mM酸化型グルタチオン-50mMリン酸カリウム-5mMエチレンジアミン四酢酸溶液(pH10.5)である。

【0065】その後これを約50倍に濃縮し、溶媒を200mM塩化ナトリウム-5mMリン酸カリウム緩衝液(pH10)に置換した。これを巻き戻りp D 3- β 溶液と呼ぶ。

【0066】<4>自己触媒的活性化

巻き戻りp D 3- β 溶液をpH4前後、200mM塩化ナトリウムに溶液に曝し、30°Cでインキュベートさせた。この結果、分子量約41kDaのp D 3- β は自己触媒的に分子量約30kDaのD 3- β に変換した。このD 3- β のc 30分解の比活性は15U/mgで、発芽ダイズ子葉より精製した天然型D 3- β の約1/5であった。

【0067】なお、このc 30分解活性の単位である1Uとは、1 μ gのc 30を1分間に消失させる活性である。本比活性測定は、基質c 30の濃度が0.5mg/m l、反応溶液組成が50mM 酢酸ナトリウム(pH4.0)、0.2M NaCl、2mM アジ化ナトリウム、10mM 2-メルカブトエタノールとなるよう酢酸溶液を添加し、30°Cで18時間反応させた反応液の還元SDS-PAGEを行い、ゲルをクマジーブリリアントブルーで染色し、デンシトメーターでc 30のバンドの減少を定量することにより行った。基質であるc 30とは、発芽過程におけるダイズ貯蔵タンパク質7Sグロブリンの分子量30kDaの限定分解フラグメントであり、次のようにして精製した。すなわち発芽7日目のダイズ子葉抽出物から等電点沈澱で夾雜タンパク質を除いたものを、レクチンリガンドアフィニティー担体Con A Sepharose(ファルマシア社製)に供し、吸着画分を溶出させたものである。なお、この活性測定法は、本発明者らにより既に出版されている特願平6-294548号明細書に詳述されている。

【0068】このようにして達成された生産量は、c 30分解活性に換算すると、天然型D 3として約50mg/1培地であり、これは発芽10日目のダイズ子葉約50kgより精製した量に匹敵する。

【0069】なお、p GEMp β は大腸菌J M 109株に形質転換され、その保持菌はE s c h e r i c h i a c o l i A J 13126として、工業技術院生命工学工業技術研究所に平成7年7月3日に寄託されている(受託番号: F E R M P-15021)。

【0070】[実施例4 D 3- β の酵母での分泌生産] 実施例2で得られたD 3- β c DNAの一部を酵

13

母で機能する発現ベクターに組み込み、この発現プラスミドを保持する形質転換体を培養し、培地中にD3活性を検出した。

【0071】このときベクターに組み込んだDNAは、配列表の配列番号2のアミノ酸配列番号-132～241位をコードする部分（すなわち塩基番号48～1166部分）に終始コドンTGA（すなわち配列番号3の塩基番号1054～1056部分）を連結したものである。これは、D3-β-cDNAの塩基配列より推定される全一時構造から、C末端側のアミノ酸配列番号242位のCys～332位のA1a部分を除いたものである。すなわち、D3-βのシグナルペプチドの下流に実施例3で発現させた配列を連結したものであり、該シグナルペプチドは配列表の配列番号2のアミノ酸配列番号-132～-108位をコードする部分（すなわち塩基番号48～122部分）にあたる。

【0072】酵母を用いてD3を培地に分泌生産させると、D3の折り畳みが行われるためD3活性体が得られ、またD3活性の宿主への影響が軽減できる。また、培地中には宿主由来の夾雜タンパク質が少ないため、精製が必要な場合には、操作が簡便になる等の利点がある。

【0073】以下に酵母を用いた本実施例について、<1>分泌発現プラスミドの構築、<2>該分泌発現プラスミドを用いた酵母の形質転換体の作製、<3>形質転換体の培養と活性測定について分説する。

【0074】<1>分泌発現プラスミドの構築

図3に示すように、実施例2に記載のD3-β-cDNAを組み込んだpUC18-aからcDNA配列を制限酵素KpnI、BamHIで切り出した。まず、この断片を市販の発現ベクターpYES2（インビトロゲン社製）のマルチクローニングサイトのKpnI、BamH I間に挿入し、プラスミドpYESD3(pA)を構築した。

【0075】なお、発現ベクターpYES2は、マルチクローニングサイトの上流にガラクトースで発現誘導されるGAL1プロモーターを、下流にCYC1ターミネーターを有し、マーカーとしてURA3遺伝子を有する。

【0076】次に、図4に示すように、上述のpYESD3(pA)からポリA配列を除くために、D3-β-cDNAの3'非翻訳領域のBgIIIより3'末端側の配列を除去した。すなわち、pYESD3(pA)のPmaCI、BamHI間に、pYESD3(pA)のPmaCI、BgIII断片を挿入し、pYESD3を構築した。このpYESD3は、D3-βの全長、すなわち配列表の配列番号2記載のアミノ酸残基番号-132位のMet～332位のA1aを発現するよう構築されている。

【0077】さらに、配列表の配列番号2記載のアミノ

14

酸配列番号242位のCys～332位のA1a部分を除いたもの、すなわち-132位のMet～241位のVai部分を発現するようなプラスミドを構築した。すなわち、図4に示すように、pYESD3(pA)のPmaCI、BamHI間に、実施例3記載のプラスミドpGEMpβのPmaCI、BamHI断片を挿入し、プラスミドpYESD3(ΔTerro)を構築した。

【0078】<2>発現プラスミドを用いた酵母の形質転換体の作製

10 酪酸リチウム法に従った、酵母細胞形質転換キットであるイーストメーカー（クロントック社製）を用いて、上述のプラスミドを、酵母INVSC1株（インビトロゲン社）に導入し、ウラシル非要求性の形質を有する形質転換体を得た。

【0079】<3>形質転換体の培養と活性測定

上記の形質転換体、すなわち、ベクターpYES2、pYESD3、pYESD3(ΔTerro)をそれぞれ有する酵母INVSC1株を、カゼイン重層寒天培地にて培養し、プロテアーゼ分泌生産の指標となるハローの形成を調べた。

【0080】このカゼイン重層寒天培地の組成は、下層はロイシン、ヒスチジン、トリプトファン、および糖源を含むSD寒天培地であり、上層は0.5% (w/v) のカゼインを含む寒天培地である。

【0081】糖源としてガラクトースを添加した場合に、pYESD3(ΔTerro)を有する酵母INVSC1株のみが、ハローを形成した。

30 【0082】次に、このpYESD3(ΔTerro)形質転換体のハロー形性能が、分泌されたチオールプロテアーゼD3によることの確認を行った。すなわち、上記の形質転換体をガラクトースとカザミノ酸を含むSD培地にて培養し、その培地について、合成ペプチド基質を用いてD3活性を、SDS-PAGEを用いてD3遺伝子産物を検出した。

【0083】D3活性の検出は、発現プラスミドpYESD3(ΔTerro)を保持する形質転換体である酵母INVSC1株を培養した培地125μlに、基質としてベンジルオキシカルボニル-Phe-Arg-4-Mチルクマリル-7-アミド（以下、Z-FR-MCAと略す）を用い、遊離する7-アミノ-4-メチルクマリン（以下、AMCと略す）の蛍光強度を測定することにより行った。AMCの励起波長は370nm、蛍光波長は460nmである。

【0084】このD3活性測定操作は、既知の方法（A. J. Barrett & H. Kirschke, "Methods Enzymol.", 80, 535(1981)）により、以下のようにして行った。すなわち、基質濃度が10μM、反応溶液組成が50mM 酪酸ナトリウム(pH4.0)-0.2M NaCl-2mM アジ化ナトリウム-10mM 2-メルカプトエタノールとなるように培地に添加し、37°Cで5分間反

15

応させた後に、阻害剤である(L-3トランスカルボキシオキシラン-2-カルボニル)-L-ロイシン-アグマチン(以下、E-64と略記する)を添加して反応を停止し、AMCの蛍光強度を測定した。対照として、上記基質に阻害剤E-64を加えたものを用い、これと上記貯能液を培地に添加し、同様の操作を行った。

【0085】本チオールプロテアーゼD3が合成ペプチド基質であるZ-FR-MCAをよく分解してAMCを遊離させること、およびE-64によって阻害を受けることは、本発明者らによって確認されている(特願平6-294548号明細書)。

【0086】また、対照としてD3遺伝子をもたないベクターpYES2を保持する形質転換体についても、同様な活性測定操作を行った。

【0087】AMCの蛍光強度の測定結果を表1に示す。なお、この条件において測定した1μM AMCの蛍光強度は0.64であった。

【0088】

【表1】

基 質	プラスミドを保持した形質転換体	
	pYES2	pYESD3 (ΔTer o)
+Z-FR-MCA	0.449	0.759
+Z-FR-MCA +E-64	0.448	0.296

表1から明らかなように、pYESD3(ΔTer o)形質転換体を培養した培地中には、E-64で阻害される、Z-FR-MCAを分解してAMCを生成させる活性が検出された。pYES2形質転換体を培養した倍地中には本活性は検出されなかった。

【0089】また、pYESD3(ΔTer o)形質転換体を培養した培地を濃縮してSDS-PAGEに供し銀染色を行った。このSDS-PAGE像を図5に示す。同図中、レーン1は分子量マーカーであり、低分子量からそれぞれ、14.5kDa、21.5kDa、31kDa、45kDa、66kDaである。レーン2および3は、pYES2形質転換体を培養した培地であり、各レーンにはそれぞれ培地8μl、約1ml分を添加した。レーン4および5は、pYESD3(ΔTer o)形質転換体を培養した培地のそれぞれ8μl、約2mlを添加した。

【0090】図5から明らかなように、pYESD3(ΔTer o)形質転換体を培養した培地から分子量約40kDaのpD3-βと考えられる幅の広いバンド、および30kDaのD3-βと推定されるバンドが検出された。一方、pYES2形質転換体を培養した培地には、このバンドは検出されなかった。さらに、pYESD3(ΔTer o)形質転換体を培養した培地では、培地中の夾雜タンパク質のバンドが消失しているのが観察

16

された。

【0091】上述の分析結果から、pYESD3(ΔTer o)形質転換体のカゼイン重層寒天培地上でのハローフームは、本形質転換体がプロ型D3を分泌生産し、これがD3活性体へと変換されることに起因するためと考えられる。

【0092】なお、pYESD3(ΔTer o)は酵母INVSC1株に形質転換され、その保持菌はSaccharomyces cerevisiae AJ14703として、工業技術院生命工学工業技術研究所に平成7年11月28日に寄託されている(受託番号:FERM P-15310)。

【0093】なお、本発明においては、発現ベクターに組み込まれるD3遺伝子の配列は上記実施例3および4に限られるものではない。すなわち、配列表の配列番号1または2、あるいは3に示す塩基配列またはその部分配列を表すものであって、それらがコードするアミノ酸配列が発芽ダイズ子葉に由来する天然型チオールプロテアーゼの活性特性を発揮し得るチオールプロテアーゼをコードする遺伝子であれば、本発明の範囲内に含まれる。

【0094】また、配列表の配列番号3の塩基番号331~1053で表される塩基配列はD3-β成熟体をコードする遺伝子と考えられることから、これを部分配列として含む塩基配列であって、組み換え遺伝子手法により発芽ダイズ子葉に由来する天然型チオールプロテアーゼと同効の酵素活性を発揮し得るものを作り得る塩基配列からなる遺伝子もまた、本発明範囲内に含まれる。

【0095】さらに、これらD3遺伝子を発現生産させるための配列、すなわち高発現のための異種融合タンパク質の配列や、分泌発現のためのシグナル配列が付加された配列も、本発明範囲内に含まれる。

【0096】【発明の効果】以上詳述したように、本発明によれば、ダイズ貯蔵タンパク質をアミノ酸または低分子ペプチドにまで分解し得る新規チオールプロテアーゼD3のアミノ酸配列をコードするcDNAが得られる。これらcDNAの構造情報をもとに、大腸菌や酵母などで組換え体チオールプロテアーゼを量産できる。また、上記cDNAを用いれば、実施例に詳述した以外の宿主でも、チオールプロテアーゼD3を生産することが可能である。

【0097】

【配列表】

17

配列の長さ：1807

配列の型：核酸

鎮の数：二本鎮

トポロジー：直鎖状

配列の種類:cDNA to mRNA

起源

1

*生物名: Glycine max

配列の特徴

特徴を表わす記号：CDS

存在位置：68..1459

特徴を決定した方法 : P

配列：

TCACCTTTCT CCTAAAACAT CTCCAGTGGG TTCTTCTCT TCTATTACAA	50
CAACGTGTC CAAGACC ATG ACC ATG GCC ATG GCC ACG ATC CTG CTC	97
Met Thr Met Ala Met Ala Thr Ile Leu Leu	
-130 -125	
CTG TTC ACG GTC TTC GCG GTT TCA TCG GCC CTA GAC ATG TCG ATA ATC	145
Leu Phe Thr Val Phe Ala Val Ser Ser Ala Leu Asp Met Ser Ile Ile	
-120 -115 -110	
TCG TAC GAC AAC GCC CAC GCC GCC ACG TCG CGC AGC GAC GAG GAG CTG	193
Ser Tyr Asp Asn Ala His Ala Ala Thr Ser Arg Ser Asp Glu Glu Leu	
-105 -100 -95 -90	
ATG TCC ATG TAC GAG CAG TGG CTG GTG AAG CAC GGC AAG GTG TAC AAC	241
Met Ser Met Tyr Glu Gln Trp Leu Val Lys His Gly Lys Val Tyr Asn	
-85 -80 -75	
GCG CTC GGG GAG AAG GAG AAG AGG TTC CAG ATC TTC AAG GAC AAC CTG	289
Ala Leu Gly Glu Lys Glu Lys Arg Phe Gln Ile Phe Lys Asp Asn Leu	
-70 -65 -60	
CGA TTC ATC GAC GAC CAC AAC TCC CAA GAG GAC CGA ACC TAC AAG CTC	337
Arg Phe Ile Asp Asp His Asn Ser Gln Glu Asp Arg Thr Tyr Lys Leu	
-55 -50 -45	
GGA CTG AAC CGG TTC GCG GAT CTC ACC AAC GAG GAG TAC AGG GCC AAG	385
Gly Leu Asn Arg Phe Ala Asp Leu Thr Asn Glu Glu Tyr Arg Ala Lys	
-40 -35 -30	
TAC TTG GGA ACC AAG ATC GAT CCC AAC CGG AGG CTC CGC AAG ACC CGG	433
Tyr Leu Gly Thr Lys Ile Asp Pro Asn Arg Arg Leu Gly Lys Thr Pro	
-25 -20 -15 -10	
AGC AAC CGA TAC GCG CCA CGT GTC GGC GAC AAA CTA CCT GAA TCG GTT	481
Ser Asn Arg Tyr Ala Pro Arg Val Gly Asp Lys Leu Pro Glu Ser Val	
-5 1 5	
GAT TGG AGG AAG GAA GGT GCT GTT CCT CCA GTC AAA GAC CAA GGA GGC	529
Asp Trp Arg Lys Glu Gly Ala Val Pro Pro Val Lys Asp Gln Gly Gly	
10 15 20	
TGT GGG AGC TGT TGG GCA TTC TCA GCA ATC GGT GCA GTG GAA GGA ATA	577
Cys Gly Ser Cys Trp Ala Phe Ser Ala Ile Gly Ala Val Glu Gly Ile	
25 30 35	
AAT AAG ATA GTG ACA GGT GAA CTG ATT TCG TTA TCA GAA CAA GAA TTG	625
Asn Lys Ile Val Thr Gly Glu Leu Ile Ser Leu Ser Glu Gln Glu Leu	
40 45 50 55	
GTG GAT TGT GAT ACA GGA TAT AAC GAA GGA TGC AAT GGA GGA CTT ATG	673
Val Asp Cys Asp Thr Gly Tyr Asn Glu Gly Cys Asn Gly Gly Leu Met	
60 65 70	
GAC TAT GCA TTT GAG TTC ATT ATC AAC AAT GGC GGC ATT GAT TCT GAA	721
Asp Tyr Ala Phe Glu Phe Ile Ile Asn Asn Gly Gly Ile Asp Ser Glu	
75 80 85	

1 9

GAG GAT TAC CCT TAC CGT GGT GTT GAT GGT AGA TGT GAC ACA TAT AGG		769	
Glu Asp Tyr Pro Tyr Arg Gly Val Asp Gly Arg Cys Asp Thr Tyr Arg			
90	95	100	
AAA AAT GCT AAG GTT GTT TCT ATT GAT GAC TAC GAA GAT GTT CCT GCC		817	
Lys Asn Ala Lys Val Val Ser Ile Asp Asp Tyr Glu Asp Val Pro Ala			
105	110	115	
TAT GAT GAG TTA GCT TTG AAA AAG GCT GTT GCA AAT CAG CCC GTC AGT		865	
Tyr Asp Glu Leu Ala Leu Lys Ala Val Ala Asn Gln Pro Val Ser			
120	125	130	135
GTA GCT ATT GAA GGA GGG GGC AGG GAA TTT CAA TTA TAT GTA TCT GGT		913	
Val Ala Ile Glu Gly Gly Arg Glu Phe Gln Leu Tyr Val Ser Gly			
140	145	150	
GTA TTC ACT GGG AGA TGT GGC ACA GCA CTA GAT CAT GGT GTC GTG GCT		961	
Val Phe Thr Gly Arg Cys Gly Thr Ala Leu Asp His Gly Val Val Ala			
155	160	165	
GTT GGG TAT GGT ACA GCT AAT GGT CAT GAT TAT TGG ATC GTA AGG AAT		1009	
Val Gly Tyr Gly Thr Ala Asn Gly His Asp Tyr Trp Ile Val Arg Asn			
170	175	180	
TCA TGG GGT CCT AGC TGG GGA GAA GAT GGC TAC ATC AGG TTA GAA AGA		1057	
Ser Trp Gly Pro Ser Trp Gly Glu Asp Gly Tyr Ile Arg Leu Glu Arg			
185	190	195	
AAT CTT GCT AAC AGC AGA TCA GGC AAG TGT GGA ATT GCA ATT GAG CCA		1105	
Asn Leu Ala Asn Ser Arg Ser Gly Lys Cys Gly Ile Ala Ile Glu Pro			
200	205	210	215
TCT TAT CCC CTT AAG AAT GGT CCA AAT CCA CCT AAT CCT GGA CCA TCA		1153	
Ser Tyr Pro Leu Lys Asn Gly Pro Asn Pro Pro Asn Pro Gly Pro Ser			
220	225	230	
CCC CCT TCA CCT GTG AAG CCA CCA AAT GTC TGT GAC AAC TAC TAC AGC		1201	
Pro Pro Ser Pro Val Lys Pro Pro Asn Val Cys Asp Asn Tyr Tyr Ser			
235	240	245	
TGT GCT GAT AGT GCT ACT TGT TGC TGT ATT TTC GAG TTC GGA AAT GCT		1249	
Cys Ala Asp Ser Ala Thr Cys Cys Cys Ile Phe Glu Phe Gly Asn Ala			
250	255	260	
TGC TTT GAG TGG GGT TGC TGT CCT CTT GAG GGT GCT ACC TGC TGT GAT		1297	
Cys Phe Glu Trp Gly Cys Cys Pro Leu Glu Gly Ala Thr Cys Cys Asp			
265	270	275	
GAC CAC TAC AGT TGC TGC CCT AAC GAC TAT CCC ATC TGC AAC ACT TAT		1345	
Asp His Tyr Ser Cys Cys Pro Asn Asp Tyr Pro Ile Cys Asn Thr Tyr			
280	285	290	295
GCT GGA ACT TGT CTC AGG AGC AAG AAC AAC CCA TTC GGA GTG AAG GCA		1393	
Ala Gly Thr Cys Leu Arg Ser Lys Asn Asn Pro Phe Gly Val Lys Ala			
300	305	310	
TTG AGG CGT ACT CCG GCT AAA CCC CAT TGG ACC TTT GGG CGT AAG AAC		1441	
Leu Arg Arg Thr Pro Ala Lys Pro His Trp Thr Phe Gly Arg Lys Asn			
315	320	325	
AAG GTC AGC AGT GCT TAAGCAGATA AAGGAAATGT GAACGTGCAG GAAGGAGCCA		1496	
Lys Val Ser Ser Ala			
330			
ATGATGAAGT AAATGAAAGA CGAAGAATTA AAACACCAGC TGATACACTT CCCCTTAC 1556			
TTCACATAAA TTTCTGTG CCATTTATT ATGATACAAA TTGATACTAT TCTGTATCAA 1616			

21

ATTCAGTCA GATATTGGCT GTGTACATAG GGTTATATTA TTGATGCACT GTATTTGTAT 1676
 TCAAATTCCA ATTCCAAATG GCGAACATAT TTGATTGCTG TTAGTTCCC AGTTTATGTA 1736
 GACTATGTA CAATGCAGCA TTATTTGGAT TTGCCGATA TTCTGGCTAT TTCATTTGCT 1796
 TAAAAAAAAA A 1807

【0098】配列番号：2

配列の長さ：1849

配列の型：核酸

鎖の数：二本鎖

トポロジー：直鎖状

配列の種類：cDNA to mRNA

*起源

生物名：Glycine max

配列の特徴

特徴を表わす記号：CDS

存在位置：48..1442

* 10 特徴を決定した方法：P

配列：

TAAACATCTC AATGGGTCT TCTTCTATTA CAACATGCC CGCAACC	47
ATG ACC ATG GCT GCG ATC GTG CTC CTG TTC ACG GTC TTT GCC GTT TCC	95
Met Thr Met Ala Ala Ile Val Leu Leu Phe Thr Val Phe Ala Val Ser	
-130 -125 -120	
TCC GCC CTA GAC ATG TCG ATA ATC TCG TAC GAC AGC GCC CAC GCG GAC	143
Ser Ala Leu Asp Met Ser Ile Ile Ser Tyr Asp Ser Ala His Ala Asp	
-115 -110 -105	
AAG GCC GCC ACG TTG CGC ACC GAG GAG GAG CTG ATG TCC ATG TAC GAG	191
Lys Ala Ala Thr Leu Arg Thr Glu Glu Glu Leu Met Ser Met Tyr Glu	
-100 -95 -90 -85	
CAG TGG CTC GTG AAG CAC GGG AAG GTG TAC AAC GCG CTC GGC GAG AAG	239
Gln Trp Leu Val Lys His Gly Lys Val Tyr Asn Ala Leu Gly Glu Lys	
-80 -75 -70	
GAG AAG CGC TTC CAG ATC TTC AAG GAC AAC CTG CGA TTC ATC GAC GAC	287
Glu Lys Arg Phe Gln Ile Phe Lys Asp Asn Leu Arg Phe Ile Asp Asp	
-65 -60 -55	
CAC AAC TCC GCG GAG GAC ACC TAC AAG CTC GGA CTG AAC CGG TTC	335
His Asn Ser Ala Glu Asp Arg Thr Tyr Lys Leu Gly Leu Asn Arg Phe	
-50 -45 -40	
GCT GAT CTC ACC AAC GAG GAA TAC AGG GCC AAG TAC TTG GGA ACC AAG	383
Ala Asp Leu Thr Asn Glu Glu Tyr Arg Ala Lys Tyr Leu Gly Thr Lys	
-35 -30 -25	
ATC GAT CCC AAC CGG AGG CTC GGA AAG ACC CGG AGC AAC CGC TAC GCG	431
Ile Asp Pro Asn Arg Arg Leu Gly Lys Thr Pro Ser Asn Arg Tyr Ala	
-20 -15 -10 -5	
CCA CGT GTC GGC GAC AAA TTG CCT GAT TCC GTT GAT TGG AGG AAG GAA	479
Pro Arg Val Gly Asp Lys Leu Pro Asp Ser Val Asp Trp Arg Lys Glu	
1 5 10	
GGT GCT GTT CCT CCT GTC AAA GAC CAA GGA GGC TGT GGG AGC TGT TGG	527
Gly Ala Val Pro Pro Val Lys Asp Gln Gly Gly Cys Gly Ser Cys Trp	
15 20 25	
GCA TTC TCA GCA ATC GGT GCA GTA GAA GGA ATA AAT AAG ATA GTA ACA	575
Ala Phe Ser Ala Ile Gly Ala Val Glu Gly Ile Asn Lys Ile Val Thr	
30 35 40	
GGC GAA CTG ATT TCG TTA TCA GAA CAA GAA TTG GTG GAT TGT GAT ACT	623
Gly Glu Leu Ile Ser Leu Ser Glu Gln Glu Leu Val Asp Cys Asp Thr	
45 50 55 60	
GGA TAT AAC CAA GGA TGC AAT GGA GGA CTT ATG GAC TAT GCA TTT GAG	671
Gly Tyr Asn Gln Gly Cys Asn Gly Gly Leu Met Asp Tyr Ala Phe Glu	

23

65	70	75	
TTC ATA ATC AAC AAT GGC GGC ATT GAT TCT GAT GAG GAT TAC CCA TAC			719
Phe Ile Ile Asn Asn Gly Gly Ile Asp Ser Asp Glu Asp Tyr Pro Tyr			
80	85	90	
CGT GGT GTT GAT GGT AGA TGC GAC ACA TAT AGG AAA AAT GCT AAA GTC			767
Arg Gly Val Asp Gly Arg Cys Asp Thr Tyr Arg Lys Asn Ala Lys Val			
95	100	105	
GTT TCT ATT GAT GAC TAC GAA GAT GTT CCT GCC TAT GAT GAG TTA GCC			815
Val Ser Ile Asp Asp Tyr Glu Asp Val Pro Ala Tyr Asp Glu Leu Ala			
110	115	120	
TTG AAA AAG GCC GTT GCA AAT CAG CCC GTG AGC GTT GCT ATT GAA GGA			863
Leu Lys Lys Ala Val Ala Asn Gln Pro Val Ser Val Ala Ile Glu Gly			
125	130	135	140
GGG GGC AGG GAA TTT CAA TTA TAT GTA TCT GGT GTA TTC ACG GGG AGA			911
Gly Gly Arg Glu Phe Gln Leu Tyr Val Ser Gly Val Phe Thr Gly Arg			
145	150	155	
TGT GGC ACA GCA CTA GAT CAT GGT GTC GTG GCT GTT GGG TAT GGA ACA			959
Cys Gly Thr Ala Leu Asp His Gly Val Val Ala Val Gly Tyr Gly Thr			
160	165	170	
GCT AAA GGT CAT GAT TAT TGG ATC GTA AGG AAT TCA TGG GGT TCT AGC			1007
Ala Lys Gly His Asp Tyr Trp Ile Val Arg Asn Ser Trp Gly Ser Ser			
175	180	185	
TGG GGA GAG GAT GGC TAC ATC AGA TTA GAA AGA AAT CTT GCT AAC AGC			1055
Trp Gly Glu Asp Gly Tyr Ile Arg Leu Glu Arg Asn Leu Ala Asn Ser			
190	195	200	
AGA TCA GGC AAG TGT GGA ATT GCA ATT GAG CCA TCT TAT CCC CTT AAG			1103
Arg Ser Gly Lys Cys Gly Ile Ala Ile Glu Pro Ser Tyr Pro Leu Lys			
205	210	215	220
AAT GGT CCA AAT CCC CCT AAT CCT GGA CCA TCA CCC CCT TCA CCT GTG			1151
Asn Gly Pro Asn Pro Pro Asn Pro Gly Pro Ser Pro Pro Ser Pro Val			
225	230	235	
AAG CCG CCA AAT GTC TGT GAC AAC TAC TAC AGC TGT GCT GAT AGT GCT			1199
Lys Pro Pro Asn Val Cys Asp Asn Tyr Tyr Ser Cys Ala Asp Ser Ala			
240	245	250	
ACT TGT TGC TGT ATT TTT GAG TTC GGA AAT GCT TGC TTC GAG TGG GGT			1247
Thr Cys Cys Cys Ile Phe Glu Phe Gly Asn Ala Cys Phe Glu Trp Gly			
255	260	265	
TGC TGT CCT CTT GAG GGT GCT TCC TGC TGT GAT GAC CAC TAC ACT TGC			1295
Cys Cys Pro Leu Glu Gly Ala Ser Cys Cys Asp Asp His Tyr Ser Cys			
270	275	280	
TGC CCT GCA GAC TAT CCC ATC TGC AAC ACT TAC GCT GGA ACT TGT CTC			1343
Cys Pro Ala Asp Tyr Pro Ile Cys Asn Thr Tyr Ala Gly Thr Cys Leu			
285	290	295	300
AGG AGC AAG AAC AAC CCC TTT GGA GTG AAG GCA TTA AGG CGT ACT CCA			1391
Arg Ser Lys Asn Asn Pro Phe Gly Val Lys Ala Leu Arg Arg Thr Pro			
305	310	315	
GCG AAA CCC CAT TGG ACC TTC GGA CGT AAG AAC AAG GTC AGC AGT GCT			1439
Ala Lys Pro His Trp Thr Phe Gly Arg Lys Asn Lys Val Ser Ser Ala			
320	325	330	
TAAGCAGTTT AAGGAAATGT GAACGTGCAG AAGGAGCCAA TGATGAAGTA AAAGGGAAGA	1499		

25

CGAAGGACTA AAGCACCAGC TGGATGCACT TCCCCCTCT TCACATAATG CTGCACGTAA 1559
 CATAAACATC TTTACCATTT TATTATGATA CAAAATAGAT AATGTTCTGT ATCAAATTTC 1619
 AGTCAGACAT TGGCTCTGTA CATAGGGTTA TATTATGAT GCACGTGATT TGTTTCAAA 1679
 TTCCAATGG CGAACATATT TGATTGCTGT TAGTTTCAA GGTTCTGTAG ATTATGTGAC 1739
 AATGCAGCAG CAATATTGG ATTTGCCA TATTACATG CTATTTAAT TTGCTTACAA 1799
 TGTGAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 1849

26

*配列の特徴

特徴を表わす記号 : CDS

存在位置 : 1..1056

10 特徴を決定した方法 : P

特徴を表わす記号 : insertion seq

存在位置 : 1..9

特徴を決定した方法 : E

【0099】配列番号 : 3

配列の長さ : 1056

配列の型 : 核酸

鎖の数 : 二本鎖

トポロジー : 直鎖状

配列の種類 : cDNA to mRNA

起源

生物名 : Glycine max

*

配列 :

ATG	GCT	AGG	TAC	GAC	AGC	GCC	CAC	GCG	GAC	AAG	GCC	GCC	ACG	TTG	CGC		48		
Met	Ala	Arg	Tyr	Asp	Ser	Ala	His	Ala	Asp	Lys	Ala	Ala	Thr	Leu	Arg				
																5	10	15	
ACC	GAG	GAG	GAG	CTG	ATG	TCC	ATG	TAC	GAG	CAG	TGG	CTC	GTG	AAG	CAC		96		
Thr	Glu	Glu	Glu	Leu	Met	Ser	Met	Tyr	Glu	Gln	Trp	Leu	Val	Lys	His				
																20	25	30	
GGG	AAG	GTG	TAC	AAC	GCG	CTC	GGC	GAG	AAG	GAG	AAG	CGC	TTC	CAG	ATC		144		
Gly	Lys	Val	Tyr	Asn	Ala	Leu	Gly	Glu	Lys	Glu	Lys	Arg	Phe	Gln	Ile				
																35	40	45	
TTC	AAG	GAC	AAC	CTG	CGA	TTC	ATC	GAC	GAC	CAC	AAC	TCC	GCG	GAG	GAC		192		
Phe	Lys	Asp	Asn	Leu	Arg	Phe	Ile	Asp	Asp	His	Asn	Ser	Ala	Glu	Asp				
																50	55	60	
CGA	ACC	TAC	AAG	CTC	GGG	CTG	AAC	CGG	TTC	GCT	GAT	CTC	ACC	AAC	GAG		240		
Arg	Thr	Tyr	Lys	Leu	Gly	Leu	Asn	Arg	Phe	Ala	Asp	Leu	Thr	Asn	Glu				
																65	70	75	80
GAA	TAC	AGG	GCC	AAG	TAC	TTG	GGA	ACC	AAG	ATC	GAT	CCC	AAC	CGG	AGG		288		
Glu	Tyr	Arg	Ala	Lys	Tyr	Leu	Gly	Thr	Lys	Ile	Asp	Pro	Asn	Arg	Arg				
																85	90	95	
CTC	GGA	AAG	ACC	CCG	AGC	AAC	CGC	TAC	GGC	CCA	CGT	GTC	GGC	GAC	AAA		336		
Leu	Gly	Lys	Thr	Pro	Ser	Asn	Arg	Tyr	Ala	Pro	Arg	Val	Gly	Asp	Lys				
																100	105	110	
TTG	CCT	GAT	TCC	GTT	GAT	TGG	AGG	AAG	GAA	GGT	GCT	GTT	CCT	GTC			384		
Leu	Pro	Asp	Ser	Val	Asp	Trp	Arg	Lys	Glu	Gly	Ala	Val	Pro	Pro	Val				
																115	120	125	
AAA	GAC	CAA	GGA	GGC	TGT	GGG	AGC	TGT	TGG	GCA	TTC	TCA	GCA	ATC	GGT		432		
Lys	Asp	Gln	Gly	Gly	Cys	Gly	Ser	Cys	Trp	Ala	Phe	Ser	Ala	Ile	Gly				
																130	135	140	
GCA	GTA	GAA	GGA	ATA	AAT	AAG	ATA	GTA	ACA	GGC	GAA	CTG	ATT	TCG	TTA		480		
Ala	Val	Glu	Gly	Ile	Asn	Lys	Ile	Val	Thr	Gly	Glu	Leu	Ile	Ser	Leu				
																145	150	155	160
TCA	GAA	CAA	GAA	TTG	GTG	GAT	TGT	GAT	ACT	GGG	TAT	AAC	CAA	GGA	TGC		528		
Ser	Glu	Gln	Glu	Leu	Val	Asp	Cys	Asp	Thr	Gly	Tyr	Asn	Gln	Gly	Cys				
																165	170	175	
AAT	GGA	GGA	CTT	ATG	GAC	TAT	GCA	TTT	GAG	TTC	ATA	ATC	AAC	AAT	GGC		576		
Asn	Gly	Gly	Leu	Met	Asp	Tyr	Ala	Phe	Glu	Phe	Ile	Ile	Asn	Asn	Gly				

27

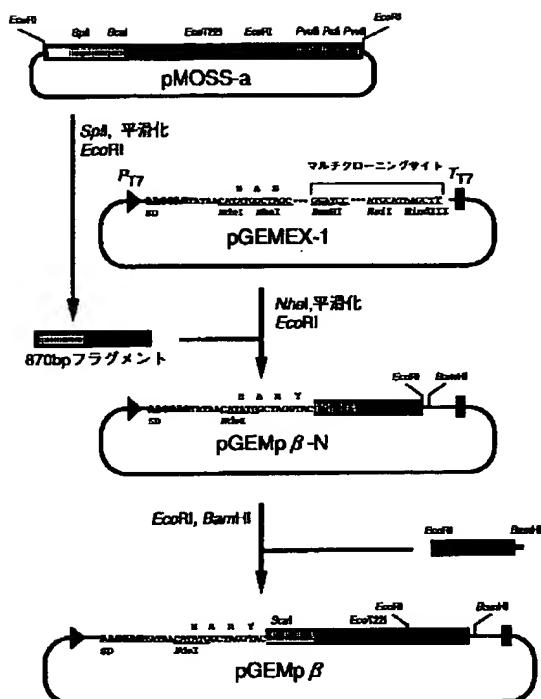
180	185	190	
GGC ATT GAT TCT GAT GAG GAT TAC CCA TAC CGT GGT GTT GAT GGT AGA			624
Gly Ile Asp Ser Asp Glu Asp Tyr Pro Tyr Arg Gly Val Asp Gly Arg			
195	200	205	
TGC GAC ACA TAT AGG AAA AAT GCT AAA GTC GTT TCT ATT GAT GAC TAC			672
Cys Asp Thr Tyr Arg Lys Asn Ala Lys Val Val Ser Ile Asp Asp Tyr			
210	215	220	
GAA GAT GTT CCT GCC TAT GAT GAG TTA GCC TTG AAA AAG GCC GTT GCA			720
Glu Asp Val Pro Ala Tyr Asp Glu Leu Ala Leu Lys Lys Ala Val Ala			
225	230	235	240
AAT CAG CCC GTG AGC GTT GCT ATT GAA GGA GGG GGC AGG GAA TTT CAA			768
Asn Gln Pro Val Ser Val Ala Ile Glu Gly Gly Arg Glu Phe Gln			
245	250	255	
TTA TAT GTA TCT GGT GTA TTC ACG GGG AGA TGT GGC ACA GCA CTA GAT			816
Leu Tyr Val Ser Gly Val Phe Thr Gly Arg Cys Gly Thr Ala Leu Asp			
260	265	270	
CAT GGT GTC GTG GCT GTT GGG TAT GGA ACA GCT AAA GGT CAT GAT TAT			864
His Gly Val Val Ala Val Gly Tyr Gly Thr Ala Lys Gly His Asp Tyr			
275	280	285	
TGG ATC GTA AGG AAT TCA TGG GGT TCT AGC TGG GGA GAG GAT GGC TAC			912
Trp Ile Val Arg Asn Ser Trp Gly Ser Ser Trp Gly Glu Asp Gly Tyr			
290	295	300	
ATC AGA TTA GAA AGA AAT CTT GCT AAC AGC AGA TCA GGC AAG TGT GGA			960
Ile Arg Leu Glu Arg Asn Leu Ala Asn Ser Arg Ser Gly Lys Cys Gly			
305	310	315	320
ATT GCA ATT GAG CCA TCT TAT CCC CTT AAG AAT GGT CCA AAT CCC CCT			1008
Ile Ala Ile Glu Pro Ser Tyr Pro Leu Lys Asn Gly Pro Asn Pro Pro			
325	330	335	
AAT CCT GGA CCA TCA CCC CCT TCA CCT GTG AAG CCG CCA AAT GTC TGA			1056
Asn Pro Gly Pro Ser Pro Pro Ser Pro Val Lys Pro Pro Asn Val			
340	345	350	

【図面の簡単な説明】

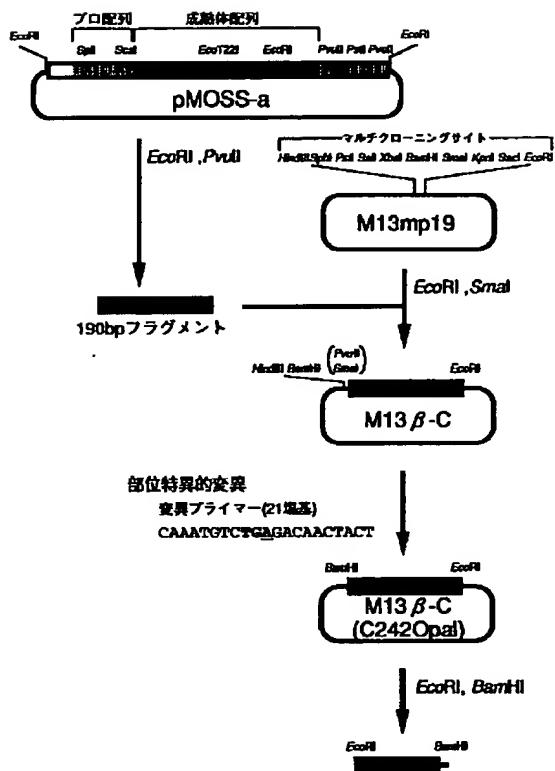
【図1】発現プラスミドpGEMp β の構築過程（部分）を示す部分図である。【図2】発現プラスミドpGEMp β の構築過程（部分）を示す部分図である。【図3】分泌発現プラスミドpYESD3（ Δ Ter \circ ）の構築過程（部分）を示す部分図である。* 【図4】分泌発現プラスミドpYESD3（ Δ Ter \circ ）の構築過程（部分）を示す部分図である。【図5】ベクターpYES2、および発現プラスミドpYESD3（ Δ Ter \circ ）を保持する形質転換体である酵母INVSC1株を培養した培地のSDS-PAGE像を示す図である。

*

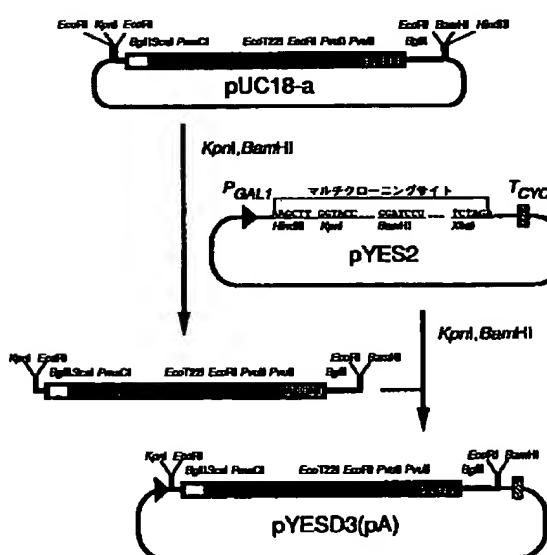
【図1】



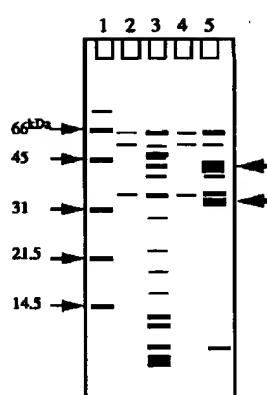
【図2】



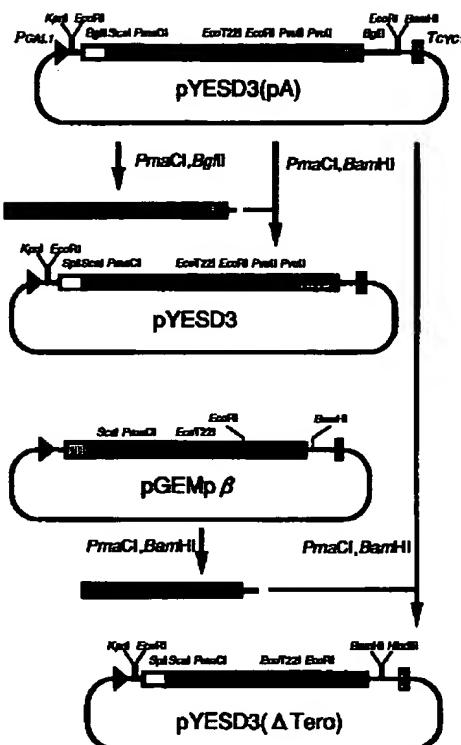
【図3】



【図5】



【図4】



フロントページの続き

(51) Int.Cl. 6	識別記号	府内整理番号	F I	技術表示箇所
C 1 2 R 1:865)				
(C 1 2 N 1/21				
C 1 2 R 1:19)				
(C 1 2 N 9/50				
C 1 2 R 1:19)				
(C 1 2 N 9/50				
C 1 2 R 1:865)				

(72)発明者 三輪 哲也
神奈川県川崎市川崎区鈴木町1番1号 味
の素株式会社食品総合研究所内

(72)発明者 柴井 博四郎
神奈川県川崎市川崎区鈴木町1番1号 味
の素株式会社中央研究所内
(72)発明者 横関 健三
神奈川県川崎市川崎区鈴木町1番1号 味
の素株式会社中央研究所内

WEST

 Generate Collection

L5: Entry 3 of 10

File: JPAB

May 13, 1997

PUB-N0: JP409121870A

DOCUMENT-IDENTIFIER: JP 09121870 A

TITLE: DNA CAPABLE OF CODING NEW THIOL PROTEASE AND PRODUCTION OF THE SAME THIOL PROTEASE USING THE DNA

PUBN-DATE: May 13, 1997

INVENTOR-INFORMATION:

NAME	COUNTRY
KAWAI, MISAKO	
ASANO, MINAO	
SUZUKI, SHUNICHI	
MIWA, TETSUYA	
SHIBAI, HIROSHIRO	
YOKOZEKI, KENZO	

INT-CL (IPC): C12 N 15/09; C12 N 1/19; C12 N 1/21; C12 N 9/50

ABSTRACT:

PROBLEM TO BE SOLVED: To obtain a new DNA, having a specific base sequence and capable of coding a new thiol protease derived from a germinated soybean cotyledon and mass-producing an enzyme capable of efficiently hydrolyzing a soybean protein with a high efficiency according to a gene recombination method.

SOLUTION: This new DNA has a base sequence represented by formula I or II and is capable of coding a new thiol protease derived from a germinated soybean cotyledon and directly mass-producing the new thiol protease capable of efficiently hydrolyzing a soybean protein into amino acids or low-molecular peptides with a high efficiency according to a gene recombination method. The new DNA is obtained by extracting an RNA from the soybean cotyledon on the 10th day after seeding according to a conventional method, then separating an mRNA from the RNA, carrying out the reverse transcription polymerase chain reaction(RT-PCR) with a primer comprising a synthetic oligonucleotide, performing the reverse transcription, amplification and cloning.

COPYRIGHT: (C)1997,JPO

us-09-80!

Mon Jan 13 10:27:28 2003

RESULT	7						OY	163 GACGTGTCACCATAACCGAGAAGGGCAAGAGACTTGATTTCAAGATACT 222
DEFINITION	E13051	1807 bp DNA	linear	PAT 27-APR-1998			Db	228 GCAGGTGTCACCATACGCCCTCGGGAGAAGGAGTTCAGTCAGGACACC 287
ACCESSION	E13051	GNA encoding thiol protease D3-alpha.					OY	223 CGACTATATCAGGAGATGAAGCANACAGAAATACCCATTCTCATCGTTAGAT 282
VERSION	E13051.1	GI: 3251863					Db	288 TGCAGTACGACGACACACTCCAGAGGACGACCTCAAG 341
KEYWORDS		UP 1997121870-A/1.					Db	342 TGACCGTTCCGGATOTCACCAACGAGGAGTACAGGGCCAGTACTTGGGACCRAGA 401
SOURCE		Glycine max.					OY	343 AGATGTGTCGAGCACATCAATGCCAACAGAATGAGAAGGAAKATAATTCT 402
ORGANISM		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; papilloideae; Phaseoleae; Glycine.					Db	402 TGCATCCAACCGAGGCTGGCAAGCCACGACGCCAGG--TGTG 458
REFERENCE		Kawai,M., Asano,M., Suzuki,S., Miwa,T., Shiba,H. and Yozeki,K. 1 (bases 1 to 1807) DNA CAPABLE OF CODING NEW THIOL PROTEASE AND PRODUCTION OF THE SAME THIOL PROTEASE USING THE DNA					OY	403 GSGACCATCCACCTGACATCATGGATGGAGAAAGGTGATCATACCCAGTAAGT 462
AUTHORS		Patent: JP 1997121870-A 1 13-MAY-1997;					Db	459 GGACAACACTACTGATGCGTGTGGAGGAGGAGTCCTGTCCTCCAGTCAGA 518
JOURNAL		AJINOMOTO CO INC					OY	523 AGCAACAGCAACAGGAGACCTGTTAGCCTCTGACAAAGAACCTGTAGCTG-- 579
COMMENT		OS Glycine max					Db	579 ATAGATGTGAGGACTGATGTTGTTAGCTAGACGAAAGGATTTGTCGATTTGATA 638
		PN JP 1997121870-A/1 13-MAY-1997					OY	580 TSGAGAGAAGGAGGTTCTAACATGGATGGCAGTACATCGTCAATGGTTAG 639
		PD 13-MAY-1997					Db	639 CAGGATAAACAGGAGTCAATGGAGACTATGGACTATGATTCATGATPATCA 698
		PF 1995-535391					OY	640 AACATGGGCGACTGCACTGTGATGATTATCCTAACAGACCAAAGGGTAGATCA 699
		PR 30-AUG-1995 JP 95P 340399					Db	699 ACATGGGGCGATGATCTGAGAGGATTACCTTACCGTGTGTGATGGTAGTG 758
		KAWAI MISAKO, ASANO MINO, SUZUKI SHUNICHI, MIWA TETSUYA, PI					OY	700 AACCCATAAGATCACAGAGGTTAACATGACGGATATGAACTCTAAATGTCAG 759
		SHIBAI HIROSHI, YOKOZEKI KENZO					Db	759 AACATGGGAAATGGAAAGGCTGTTCTATGGAGAGATGTCAGGAGATGTC 818
		PC C12N1/5/09, C12N1/19, C12N1/21, C12N9/50, (C12N1/19, C12R1:865), CC (C12N1/21, C12R1:19), (C12N9/50, C12R1:19), (C12N9/50, C12R1:865); CC					OY	760 ATGAGAGTACAGAACTGAGACAGAGGAAAGCCTCTAACGCCCATCTGAGCAACCAA 819
		strandedness: Double;					Db	819 ATGAT-----GAGTTAGCTTGAAGAAAGGCTGTGCAAACTACCG 860
		CC topology: Linear;					OY	820 T"AGTGTCTCAATTGATGCAAAAGATTTCATTAACACGGGGATTTAGATGGAG 879
		CC hypothetical: No;					Db	861 TCTGTGAGTATGAGGGGGCAGGAAATTCTATTATATG---TATCTGTGT 917
		CC anti-sense: No;					OY	880 AAACCTGACAGTCCTATGCGATTATCACTTGTTTACTTGCGGTATGGTCAG 939
		Key FH source					Db	918 TCTCTGGAGAGTGGCACAGCACTGATGTTGCTGTTGGGTATGGPACAG 977
		FT source	1. 1807 /organism="Glycine max"				OY	940 CGGATGGTGTAGTACTGGATAGCGAAATTCTATGGAGAGATTGGGGAGATG 999
		FT 5'UTR	2. 1807 /db_xref="taxon:3847"				Db	978 CTATGTCAGTATGGATGTAAGGATCTGGGTCCTAGTGGGAGANGA 1037
		FT sig-peptide	3. 1807 /product="signal Peptide of thiol protease FT				OY	1000 GTTACATTTGGTCCAAAGAACACCGGTTAA 1030
		FT	4. 1807 /product="D3-alpha" FT CDS				Db	1038 GCTCATCAGGTAGAGAAGCTCTGCTAA 1068
		FT	5. 1807 /product="thiol protease D3-alpha" FT 3'UTR					
		FEATURES	1. 1807 /location="Qualifiers					
		source	2. 1807 /organism="Glycine max"					
		BASE COUNT	500 a 387 c 449 g 471 t					
		ORIGIN						
Query Match	14.4%	Score 167; DB 6; Length 1807;						
Best Local Similarity	53.0%	Pred. No. 1. 3e-32;						
Matches	493;	Conservative 0; Mismatches 405; Indels 33; Gaps 5;						
OY	103 CCAAGCTTACACACAGAACAGGCTCTACTGTTCAACTATGGAGGTGACATG 162							
Db	168 CCACGTCGGCAGCGAGGAGCTGATGTCATGACGAGCAGTGGCTGGTGAAC 227							

(AJIN) AJINOMOTO ...

PA

XX

DR

DR

XX

PT

New thiol protease and corresponding DNA sequence - used in the preparation of the thiol protease

XX

PS Claim 1; Page 10-12; 17pp; Japanese.

WPI: 1997-314230/29.

P-PSDB; RANI19541.

The present sequence encodes a novel thiol protease D3-alpha derived from germinated soybean cotyledon. A method for the production of thiol protease using Escherichia coli transformed with recombinant expression vectors containing a DNA sequence encoding thiol protease has been produced. The method can be used to prepare large amounts of thiol protease with high efficiency.

XX Sequence 1807 BP; 500 A; 387 C; 449 G; 471 T; 0 other;

Query Match 14.4%; Score 167; DB 18; Length 1807;

Best Local Similarity 53.0%; Pred. No. 6.4e-39;

Matches 493; Conservative 0; Mismatches 405; Indels 33; Gaps 5;

QY 103 CCAAGTTTACACAGACAAGAACAGGAGCTCAGCTTCCACTATGGAAAGAGGAGCATG 162

DB 168 CCACCTCGCCAGCACGAGGAGCTGATGTCATGTACAGGAGCAGGGCTGTTGAGCACG 227

QY 163 GACGAGCTTACCATACCAAGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 222

DB 228 GCAASGTGTAACACCGCTGGAGAACAGAGTCTGAGTCAAGTCTCAGAGTCTCAGAGCACACC 287

QY 223 CGACATATATCAGGGACATGATGCAACAGAAATCACCACATTCTCATCGTTAGGAT 282

DB 288 TGCGATCTCATGCACCAACRACCTCCAAGAGGAGCCACCTCAAGC-----TCGGAC 341

QY 283 TGAACAGTTGCTGACATCATCTCTCAAGAGTTAGCAAAAGAGACTTGCAAGCTCCCA 342

DB 342 TGAACCGGTTGGGATCTCACCAACGAGGAGTACAGGCCAGTACTTGGACCAAGA 401

QY 343 AGGATGTTGTCGACGAAATCAAATGCCAACAGAAGATAAGAGGAAACAAATTTCTT 402

DB 402 TCGATCCAAACGGAGGCTCGCAAGACCCCGAACCGATACGGCCACCG---TGTG 458

QY 403 GTGACCATCCACCTGCATCATGGATGGAGAAAAGGTGTCATCACCCAGTAAATG 462

DB 459 GCGACAACTACCTGATCGGTTGATGGAGGAGGAGGCTGTCCTCCAGTCAG 518

QY 463 ACCAAGGGGGCTGGAGGGGTTGGGGTTCTGCCACGGGAGCCATAGAACGAC 522

DB 519 ACCAAGGGGCTGGGAGCTGAGGAACTCTGTTACCTCTGTCAGAACAGACTGTAGCTG--- 579

QY 523 ATGCAATGAGGAGGACTCTGTTATCTAGAACAGAATGGGAGGATAA 578

DB 579 ATGATAGTGTGAGGAGGACTCTGTTATCTAGAACAGAATGGGAGGATA 638

RESULT 3
ID AAT4041
ID AAT4041 standard; cDNA to mRNA; 1807 BP.

XX AC AAT4041;

DT 17-SEP-1997 (first entry)

XX DE soybean thiol protease D3-alpha encoding cDNA.

XX KW Soybean; cotyledon; germinate; transformant; Escherichia coli;

XX KW recombinant expression vector; ds.

XX OS Glycine max.

XX PH Key Location/Qualifiers

FT 68..1459

FT /*tag= a

FT /product= Thiol_protease

FT 68..460

FT /*tag= b

FT 461..1459

FT /*tag= c

FT /product= D3-alpha

XX PN JP09121870-A.

XX PD 13-MAY-1997.

XX XX 28-DEC-1995; 95JP-0353931.

XX PR 30-AUG-1995; 95JP-0245279.

PR 29-DEC-1994; 94JP-0340399.

JFH 09121870 A

QY 760 ATGAGACTACAGATCAGAGACAGCAAGCTCTTAAGCGCATCCCTGACCAACCA 819

DB 819 ATGAT-----GAGTAGCTTGAAGAAGGCTGTTGCAAAATGCCCC 860

820 TTAGTGTCTCAATGATGCAAAAGATTTCATTTACACGGGGAAATTATGATGGAG 879